Remarks

Claims 1–3, 5–6, and 8–28 are pending in this application. Claims 5, 9, 12–13, and 25–27 have been amended to make editorial changes and to more specifically claim the invention. The amended claims are fully supported by the specification. No new matter has been added.

Double Patenting Rejection

Claims 1–3, 5–6, and 9–21 have been rejected for obviousness-type double patenting as being unpatentable over claims 1–26 of U.S. patent 6,865,536 in view of U.S. patent 5,054,085 (Meisel). Applicant may overcome this rejection by submitting a terminal disclaimer. However, applicant will defer submitting a terminal disclaimer until the claims are allowable apart from the obviousness-type double patenting rejection.

Section 103 Rejections

Claims 1–3 were rejected under section 103 as being unpatentable over U.S. patent 5,960,399 (Barclay) in view of Meisel. Claim 5 and 6 are rejected under section 103 as being unpatentable over Barclay in view of Meisel, and further in view of U.S. patent 6,216,104 (Moshfeghi). Claims 8–13 and 18–21 have been rejected under section 103 as being unpatentable over Barclay in view of Meisel, and further in view of U.S. patent 5,751,951 (Osborne). Claims 14–17 have been rejected under section 103 as being unpatentable over Barclay in view of Meisel, and further in view of Osborne and Moshfeghi. Reconsideration of the rejections and allowance of the claims are respectfully requested for the following reasons.

Even if the references Barclay, Meisel, Osborne, and Moshfeghi were combined, and there is no suggestion to do this, the combination still falls short of the recited invention. The combination of the references does not show or suggest each and every limitation recited in the present invention.

Not Obvious to Combine Barclay and Meisel

Claim 1 recites a client to "*store* the audio speech in one or more buffers in a *raw uncompressed audio format*, each buffer comprising a portion of the received audio speech." The invention provides for a system that retains the original analog data of the speech, therefore requiring large amounts of information to be stored. In this way, specific details, such as accents and intonations, with regards to a specific user's speech are retained. Barclay and Meisel,

considered in combination, do not show or suggest retaining the original analog data of the speech in a large quantity of data.

On page 6 of the office action, the examiner states "it would have been obvious to one having ordinary skill in the art to include a feature of storing raw uncompressed audio speech in buffers as taught by *Meisel et al.* in a client/server speech processor/recognizer of *Barclay et al.*" However, the purpose of Barclay is to provide a streaming or real-time processing system that can be operated using low bandwidth. To accomplish this, Barclay extracts and quantizes only certain features of the speech, known as cepstra, for transmission. Barclay therefore teaches away from storing and transmitting the entire raw unprocessed speech, as doing so would hinder the objectives of Barclay.

Therefore, it would not have been obvious to one possessing ordinary skill in the art to combine Barclay and Meisel. Claims 1–3 and 22, and their dependent claims should be allowable for at least this reason.

No Suggestion to Combine Barclay and Meisel

There is *no suggestion or motivation to combine* Meisel with Barclay. These references are dissimilar and are inconsistent with each another.

Barclay discusses a method of speech recognition based on the unique features of a person's speech. Such a method may be used to identify or verify who is speaking. Barclay describes a system that extracts unique features of a person's speech, known as cepstra, in order to recognize the user.

In contrast to Barclay, Meisel discusses a speech preprocessing method that reduces the uniqueness of speech between different speakers. In its abstract, Meisel states:

> Thus after the pre-processing performed by this invention, the *parameters would look much the same* for the same word *independent of speaker*. In this manner, *variations in the speech signal* caused by the physical makeup of a speaker's throat, mouth, lips, teeth, and nasal cavity would be, at least in part, *reduced by the pre-processing.*

Thus, after Meisel preprocesses some speech, there would be fewer, if any, unique characteristics of that speech upon which Barclay's method may use. When Meisel is combined with Barclay, Barclay would likely not be able to identify or verify a person.

Therefore, one having ordinary skill in the art would not combine Meisel with Barclay. Barclay's method depends on maintaining the unique characteristics of each person's speech, while Meisel removes differences in speech that distinguish different speakers. There is no motivation to combine these references. For at least this reason, claims 1–3, 5–6, and 8–21 and their dependents should be allowable.

Combination Falls Short

Even if the references were combined, and there is no suggestion to do this for the reasons discussed above, the combination still falls short of the recited invention. The combination of the references does not show or suggest each and every limitation of the recited invention.

Both Barclay and Meisel describe preprocessing and compression of data for optimal processing. In Barclay and Meisel, the preprocessing and compression involves filtering of the data, where unnecessary and insignificant data (in Barclay and Meisel's view of speech processing) is removed.

At column 4, lines 3–4 and column 5, lines 3–4, Barclay specifically describes a program that extracts certain features of digitized speech, called cepstra. In extracting only these specific features, Barclay is able to send less data for communication between the client and the server. Column 4, lines 15–16. Barclay then describes a system that quantizes said features for transmission to the server. As is known in the art, the use of quantization is motivated by a need to reduce the amount of data needed to represent a signal. In no way would Barclay teach retention of the entire analog signal, as that would result in storing and transmission of unnecessary data.

On page 6 of the office action, the examiner states that "[it] would have been obvious to one having ordinary skill in the art to include a feature of storing raw uncompressed audio speech in buffers as taught by *Meisel et al.* in a client/server speech processor/recognizer of *Barclay et al.* for a purpose of providing for analysis and collection of speech data for speech recognition to obtain optimal parameters during preprocessing." However, like Barclay, Meisel does not teach the retention of the entire analog signal, as that would result in the storing and transmission of unnecessary data. The purpose of Meisel is to reduce the variations in the speech signal by the preprocessing method. Meisel thus describes a system for preprocessing where raw

data *cannot* be buffered, and therefore only specific data is collected based on a set of nominal values. Column 5, lines 13 to 21.

However, in the present invention, unlike the prior art, the speech data is transmitted and the server stores the raw speech. Therefore, even if Barclay and Meisel were to be combined, even though there is no suggestion to do so, the combination falls short of the present invention.

Therefore, claim 1 and its dependent claims should be allowable for at least this reason.

Data Communication

In the office action, the examiner states "it is well known to buffer *data communications* both upon reception and before transmission to permit processing." However, claim 1 recites that a client *stores "the audio speech in one or more buffers* in a raw uncompressed audio format," which something very different from what the examiner states. Data communication refers to communication between two nonhuman devices, such as two computers. *Claim 1 does not refer to data communication, but storing of audio speech from for example, a human user.*

For at least this additional reason, the examiner has not shown obviousness, and claim 1 and its dependent should be allowable. Claims 9 and 11 recite similar limitations as in claim 1, and these claims and their dependents should be allowable for at least similar reasons.

No Encoding of the Received Speech

Additionally, claim 1 recites a client having a capability to "*encode* a buffer of the received audio speech before all of the audio speech is received." None of the cited references, individually or in combination, show or suggest this limitation.

Barclay does not teach or suggest encoding audio speech for transmission through a communication network, but rather describes extracting and quantizing cepstral features. Only the cepstral features are sent to a server through a communication network. Column 4, lines 3–5. As discussed above, cepstral features are *not speech*. And furthermore, the extracting or quantizing of Barclay are *not "encoding."*

Extracting features from speech is not encoding the speech. Barclay describes extracting features from speech by applying some mathematical operations on the speech to form features. Column 2, lines 3–8. These features are mathematical formations from speech without the speech content and quality. Column 2, lines 16–23. Extracting is unlike encoding because extracting does not preserve the content and quality of the speech.

Quantizing features is also not encoding speech. Barclay itself distinguishes "quantizing" from "encoding." Barclay's only use of the encoding term is at column 5, lines 61–64. There, Barclay describes encoding an end-of-speech (EOS) signal before sending it to the server. This does not show or suggest the recited invention because the *EOS signal is not speech*. Rather, the EOS signal is generated by the Barclay client after "a period of silence is encountered." Column 7, lines 23–24. As shown in figure 2A, the Barclay client waits for the user to stop speaking in step 36, and then the client (not a person) generates the EOS signal in step 38.

In all other instances in the reference, Barclay uses the term quantize (i.e., 26 times), and it is cepstral features that are quantized. Barclay never describes the cepstral features as being encoded. Moreover, as has been discussed, the cepstral features are not speech.

Barclay clearly *does not teach or suggest encoding* a buffer of the received audio speech before all of the audio speech is received. Meisel also *does not teach or suggest encoding* a buffer of the received audio speech. The references do not provide the features of benefits of the present invention.

The present invention encodes a buffer of the audio speech so that the server can evaluate the audio speech. For example, an embodiment of the invention evaluates the pronunciation accuracy of the audio speech. Such a system may be used to help speakers with a nonnative accent to learn to speak without the accent. The prior art does not show or suggest a system of the invention. For at least this additional reason, claim 1 and its dependents should be allowable. Claims 9 and 11 recite similar limitations as in claim 1, and these claims and their dependents should be allowable for at least similar reasons.

Claim 9 further recites additional limitations not shown or suggested by the prior art.

Dependent Claims

Claim 5 recites "the server further comprises two or more stored text format files, and the server selects a stored text format file to transmit to a client of the two or more clients as a result of the server's evaluation of the resultant raw speech received from the client, and *the server adjusts a processing time used to evaluate the resultant raw speech based on a value in a URL connection between the client and the server*." Nowhere does the prior show or suggest adjusting the processing time used to evaluate the resultant raw speech. For at least this reason, claim 5 should be allowable.

Claim 12 recites "the response a result of the server's evaluation of the resultant raw speech received from the client, and the server *alters a processing time used to evaluate the resultant raw speech based on a value communicated between the client and the server.*" Nowhere does the prior show or suggest adjusting the processing time used to evaluate the resultant raw speech. For at least this reason, claim 12 should be allowable.

Claim 13 recites "*based on a URL sent by the client, the server determines whether to expect speech data for processing from the client.*" Nowhere does the prior show or suggest adjusting the processing time used to evaluate the resultant raw speech. For at least this reason, claim 13 should be allowable.

Claim 25 recites "*the server evaluates the resultant raw speech received from the client based on the user objective and a value communicated to the server by URL.*" Nowhere does the prior show or suggest adjusting the processing time used to evaluate the resultant raw speech. For at least this reason, claim 25 should be allowable.

Claim 26 recites "*a processing time used to evaluate the resultant raw speech will vary based on a value communicated to the server from the client.*" Nowhere does the prior show or suggest adjusting the processing time used to evaluate the resultant raw speech. For at least this reason, claim 26 should be allowable.

Claim 27 recites "*a processing time used by the server to evaluate the resultant raw speech is alterable based on a value communicated from the client to the server.*" Nowhere does the prior show or suggest adjusting the processing time used to evaluate the resultant raw speech. For at least this reason, claim 26 should be allowable.

Conclusion

For the above reasons, applicant believes all claims now pending in this application are in condition for allowance. Applicant respectfully requests that a timely Notice of Allowance be issued in this case. If the examiner believes a telephone conference would expedite prosecution of this application, please contact the signee.

Respectfully submitted,

Aka Chan LLP

/Melvin D. Chan/

Melvin D. Chan
Reg. No. 39,626

Aka Chan LLP
900 Lafayette Street, Suite 710
Santa Clara, CA 95050
Tel: (408) 701-0035
Fax: (408) 608-1599
E-mail: mel@akachanlaw.com